Optimizing facility operation in high density data center environments

technology brief

Abstract 3

Introduction 3

Assessing facility requirements 3
- Using sizing tools to understand server and facility requirements 4
- Data center assessment 5

Component power and cooling 5
- Processor power and cooling 5
  - Processor P-state 5
  - P-state management 6
- Efficient practices for servers and enclosures 6
  - Understanding server power utilization and heat generation 6
  - Greater density with multi-core processors 7
  - Blade enclosures with integrated cooling hardware 7
  - Onboard thermal sensors and temperature management 7
  - Pooled power 7
  - High-line power efficiency 7
  - Power management and reporting of individual servers 8
  - Server virtualization using virtual machine technology 9
- Efficient practices at the rack level 9
  - Rack configuration tools 9
  - Methods for calculating power supply requirements 10
  - Power protection and management 10
  - Rack based power distribution 10
  - High-efficiency power supplies 10
  - Understanding Internal airflow 10
  - Liquid cooled racks 11

Efficient practices for facility-level power and cooling 12
- Power management and reporting 13
  - Comprehensive management applications 13
  - Power Capping 13
- Airflow distribution for high-density data centers 13
  - Raised floors 15
  - Air supply plenum 15
  - Ceiling return air plenum 17
  - Dual supply air plenums 17
  - Perforated tiles 17
  - Rack geometry 18
  - Cooling footprint 18
  - Hot and cold aisle spacing 18
  - Row configuration 18

Closely coupled cooling 19
Computer room air conditioners 19
Capacity of CRAC units 19
Placement of CRAC units 20
Discharge velocity 21
Advanced thermal management techniques 22
Thermal Assessment Services 22
Dynamic Smart Cooling 24
Summary 25
For more information 26
Call to action 26

Abstract

This paper describes issues of high processor and server density within existing data center infrastructures. It identifies methods to optimize the effectiveness of power and cooling resources in facilities that are deploying high-density equipment or that are already fully populated with high-density equipment. The intended audience for this paper includes IT managers, IT administrators, facility planners, and operations staff. Managers considering strategic planning activity such as facility upgrades or new facility planning should consult the "Data Center Cooling Strategies" technical brief available at
http://h20000.www2.hp.com/bc/docs/support/SupportManual/c01153741/c01153741.pdf.

Introduction

Existing data centers have had three major operational and fiscal constraints – power, cooling and space. As existing data centers are required to support increasingly dense configurations, power and cooling requirements can outstrip the capabilities of the data center infrastructures. In fact, the issue of space becomes moot because existing data centers are projected to run out of power before they run out of space.

HP understands that individual data center challenges like rack and processor level power and cooling cannot be viewed and managed as disconnected issues. HP has developed solutions to the fundamental issues of power and cooling at the processor, server, rack, and facility infrastructure level. HP also has developed proven management tools to provide a unified approach to managing power and cooling in data centers. This document describes issues, tools, and solutions for optimizing power and cooling at each level of the data center: processor, server, rack, and facility.

The first section of this paper describes facility assessment and how the component and server level issues affect power and cooling, while the last half of the paper describes rack-level, facility layout, and data center level approaches.

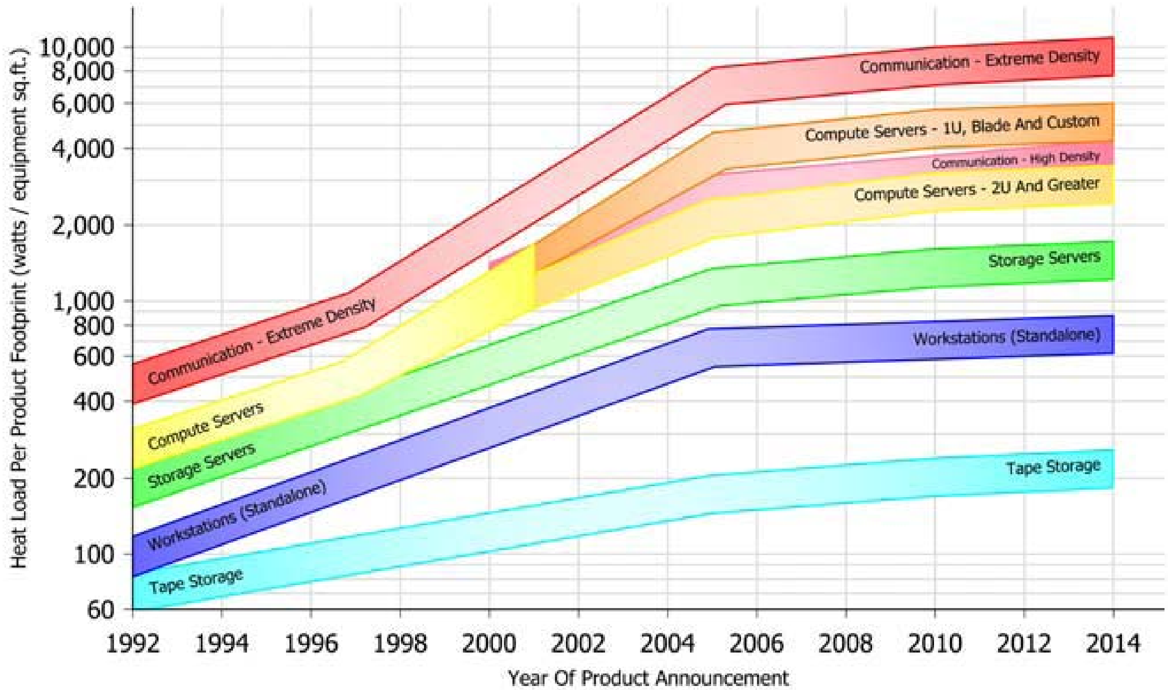
Assessing facility requirements

Data center facilities face unprecedented power and cooling requirements. For example, 2 meter rack with 42 DL360 G5 servers can theoretically use up to 29 kVA of power at peak conditions; while a 2 meter rack with 64 HP BladeSystem c-Class server blades can use up to 31 kVA at peak conditions. A power draw of 30kVA per rack occupying 30 cu. ft. of space equals a power density of 1000 watts per square foot.

However, a typical existing datacenter was built with redundant power of 3328 VA (20A single-phase) to 8640 VA (30A three-phase) and has a power/cooling density of 50 to 150 watts per square foot. As a result, capacity can be an order of magnitude too small.

Server power densities have increased up to 10 times in the last 10 years, and industry trends suggest the power density projections displayed in Figure 1.

Figure 1. Projection of product heat loads in the data center



Source: ASHRAE, Datacom Equipment Power Trends and Cooling Applications, Chapter 3, Fig 3.10 New ASHRAE updated and expanded power trend chart, © 2005 American Society of Heating, Refrigerating and Air-Conditioning Engineers, Inc. Figure is based on a fully loaded 42U rack with maximum configurations, typical power consumptions, and a heat load based on square footage of the equipment footprint.

Using sizing tools to understand server and facility requirements

Data center sizing tools are available to address both small-to-medium size business requirements and enterprise business requirements. These tools can provide power and computational workload estimates for both rack-mounted servers and server blades. Available sizing tools also address data center storage requirements and remote management implementations. These tools are typically produced by each hardware vendor. HP tools include the HP Power Calculator utility, the HP BladeSystem Power Sizer, the HP StorageWorks Sizing Tool, and the HP Systems Insight Manager (SIM) sizing tool.

The HP Power Calculator utility enables IT planners to more accurately estimate the power needs of HP ProLiant systems. The calculator provides meaningful estimates of power requirements for sizing an infrastructure. It is important to note, however, that actual power consumption of an installed system running certain applications may vary. The calculator is available as a Microsoft® Excel® downloadable, interactive catalog at http://h30099.www3.hp.com/configurator/calc/Power Calculator Catalog.xls.

The HP BladeSystem Power Sizer enables the facilities team to effectively plan for the power consumption and heat load of an HP BladeSystem. The tool is based on actual, component-level power measurements of a system stressed to maximum capability.

The HP StorageWorks Sizing Tool provides information for designing a storage infrastructure. This downloadable sizing tool applies storage design rules, licensing, and services rules to design the system.

The HP Systems Insight Manager (SIM) sizing tool assists systems administrators in building highly available, high-performance HP SIM deployments running on Microsoft Windows®, Linux®, or HP-UX operating systems. With input from quality assurance and performance testing, the tool recommends server configurations based on projected management workloads, installed management applications, and number of console users. The tool also provides performance and configuration data designed to help systems administrators optimize the performance of existing HP SIM deployments.

To see the complete offering of HP sizers and configurators, go to the HP ActiveAnswers tools web page at http://h71019.www7.hp.com/ActiveAnswers/cache/71114-0-0-0-121.html

Data center assessment

An evaluation of current workloads, facility infrastructure, and server configurations is necessary to determine existing power and cooling requirements, and to project future requirements. IT administrators and data center managers should have access to methodologies and services designed to assess existing conditions, identify problems, and indicate solutions. HP Datacenter Thermal Assessment Services[1] provide a thorough review and analysis of facility infrastructure. If risks or deficiencies are found, a qualitative and quantitative explanation is provided for each, including prioritized recommendations based on industry experience, industry standards, and engineering and operational best practices.

Component power and cooling

This section addresses efficient practices for power and cooling at a component level and the individual solutions required for each component.

Processor power and cooling

Processor power and cooling requirements depend on the per watt performance of an individual processor. The ability to manage this performance in relation to demand, at both the processor and chip levels, is crucial to avoiding power and cooling issues.

Processor P-state

The latest server processors from Intel and AMD have power state registers that are available to programmers. With the appropriate ROM firmware or operating system interface, these hardware registers can be used to switch a processor between different performance states or P-states[2]. Changing the performance state (that is, the processor frequency and voltage) enables processors to operate at different power levels. Tables 1 and 2 list P-states exposed by the Intel® Xeon™ 3.8-GHz/800-MHz and Quad-Core 2.66-GHz processors.

Table1. P-states of the Intel Xeon 3.8-GHz processor

P-state	Description	Core frequency	Approximate core voltage
Pmax	Maximum performance	3.8 GHz	1.4 VDC
Pmin	Minimum power	2.8 GHz	1.2 VDC

[1] For more information go to http://h20219.www2.hp.com/services/cache/114078-0-0-225-121.html?jumpid=reg_R1002_USEN.

[2] **P-states**—The ACPI body defines P-states as processor performance states. For Intel and AMD processors, a P-state is defined by a fixed operating frequency and voltage.

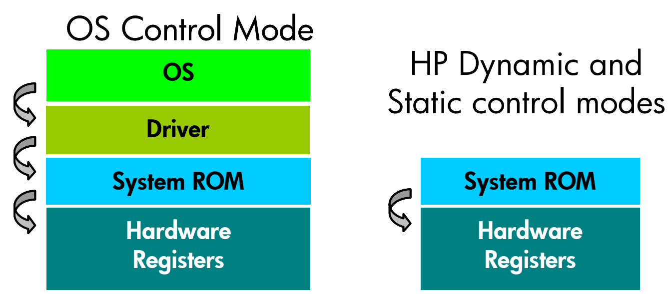
Table2. P-states of the Intel Quad-Core 2.66-GHz processor

P-state	Description	Core Frequency	Approximate Core voltage
Pmax	Maximum performance	2.66 GHz	1.2 VDC
Pmin	Minimum power	2.0 GHz	1.0 VDC

P-state management

IT administrators can control processor P-states by one of two basic methods: through the operating system (OS) with the use of a driver, or more directly through firmware in the BIOS ROM (Figure 2).

Figure 2. Methods for controlling processor P-states



An OS-based control method requires an OS upgrade and driver installation on any server where P-state management is desired. A ROM-based solution, however, provides P-state control at power-up, requires no software loading or upgrade, and can operate on systems running an OS that does not support P-state management. HP Power Regulator for ProLiant servers is an example of OS-independent, power management. This HP hardware/software implementation enables a system administrator to manage processor power consumption and system performance in either static or dynamic modes. Mare information about the Power Regulator for ProLiant servers can be found at www.hp.com/servers/power-regulator.

Efficient practices for servers and enclosures

The HP product line includes dual-processor and quad-processor server blades that can be installed in the same rack-mounted enclosure, interconnected, and easily managed. This high-density server technology lowers the operating cost per processor by reducing management expenses and the requirements for floor space.

Understanding server power utilization and heat generation

IT equipment manufacturers typically provide power and heat load information in their product specifications. HP provides a Rack/Site Installation Preparation Utility to assist customers in approximating the power and heat load per rack for facilities planning. The Site Installation Preparation Utility uses the power calculators for individual platforms so that customers can calculate the full environmental effect of racks with varying configurations and loads. This utility can be downloaded from http://h30099.www3.hp.com/configurator/calc/Site%20Preparation%20Utility.xls.

Greater density with multi-core processors

Multi-core processors take advantage of a fundamental relationship between power and frequency. Each core in a multi-core processor can run at a lower frequency, dividing among them the power normally given to a single core. The result is a performance increase over a single-core processor.

Blade enclosures with integrated cooling hardware

HP BladeSystem c-Class provides efficient cooling with the use of innovative Active Cool fan technology plus the intelligent management of the Onboard Administrator controller. HP Active Cool Fan technology optimizes airflow and reduces power draw for BladeSystem c-Class server blade configurations. The fans are controlled by the HP BladeSystem c-Class Onboard Administrator, which can increase or decrease cooling capacity based on the needs of the entire system. As a result, c-Class enclosures can accommodate full-featured servers that are 60 percent more dense than traditional rack-mount servers, while the Active Cool fans consume only 50 percent of the power typically required by fans and use 30 percent less airflow.

Onboard thermal sensors and temperature management

The inclusion of sensors within enclosures and racks, on servers and blades, and on any critical switch provides the capability to monitor changes and violations of operating thresholds. Management applications, such as HP Onboard Administrator, use that information to control server cooling and performance.

Pooled power

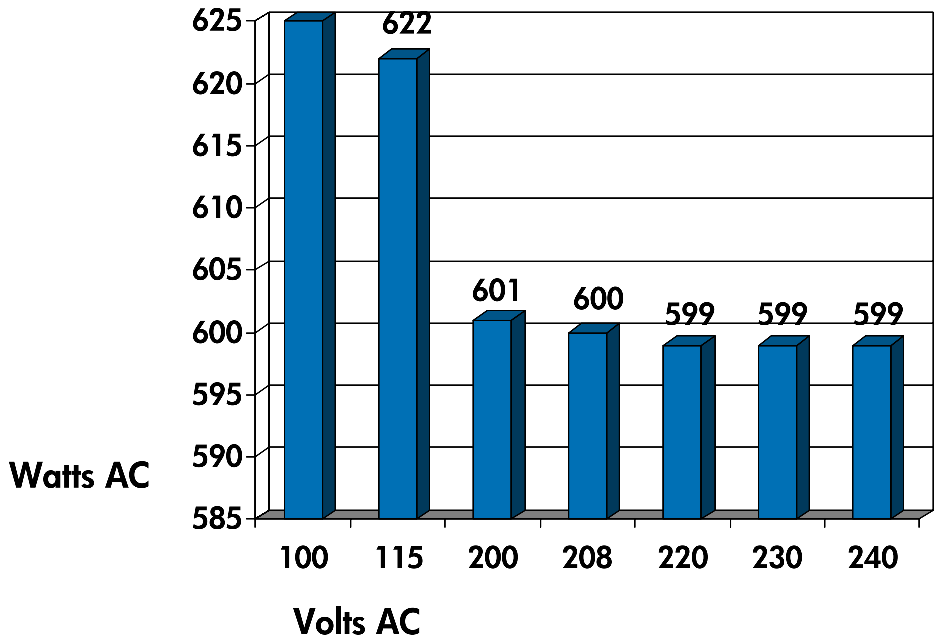
'Pooled power' improves power usage by using only the power supplies needed to match the requirements of customers' consolidated infrastructures. Since power supplies are most efficient running at higher loads, this feature keeps them working at their most efficient levels while the power supply remains fully redundant. The HP Dynamic Power Saver is a feature of the BladeSystem c-Class and p-Class blade enclosures. The HP Dynamic Power Saver runs continuously in the background, pooling power distribution to maintain system performance at higher application loads and providing power savings at lower application loads. More information on 'Pooled Power' can be found at http://h20000.www2.hp.com/bc/docs/support/SupportManual/c00816246/c00816246.pdf.

High-line power efficiency

In the Americas and other areas that follow the same commercial wiring practices, organizations have the choice between low-line power (100-120V AC) and high-line power (200-240V AC) for their servers. This is an important choice, since high-line service is the most stable, efficient, and flexible AC power for server and data operations. High-line power offers greater efficiencies than single-phase power.

Measuring the power consumption of a ProLiant DL380 G4 server demonstrates the higher efficiencies of high-line power. A test using the same system configuration, running a typical application, but using different AC supply voltages returned the power consumption data in Figure 3.

Figure 3. Power consumption of ProLiant DL380 G4 server at different AC supply levels



Therefore, a 1000-server datacenter would save approximately $25,000 per year[3] in direct and indirect power costs by using 208V instead of 115V power.

Power management and reporting of individual servers

Power management and reporting of all servers are essential to managing power and cooling in a data center environment. HP provides the HP Power Regulator for ProLiant, HP Integrated Lights Out processors (iLO and iLO 2), and HP SIM to help manage power at a server level.

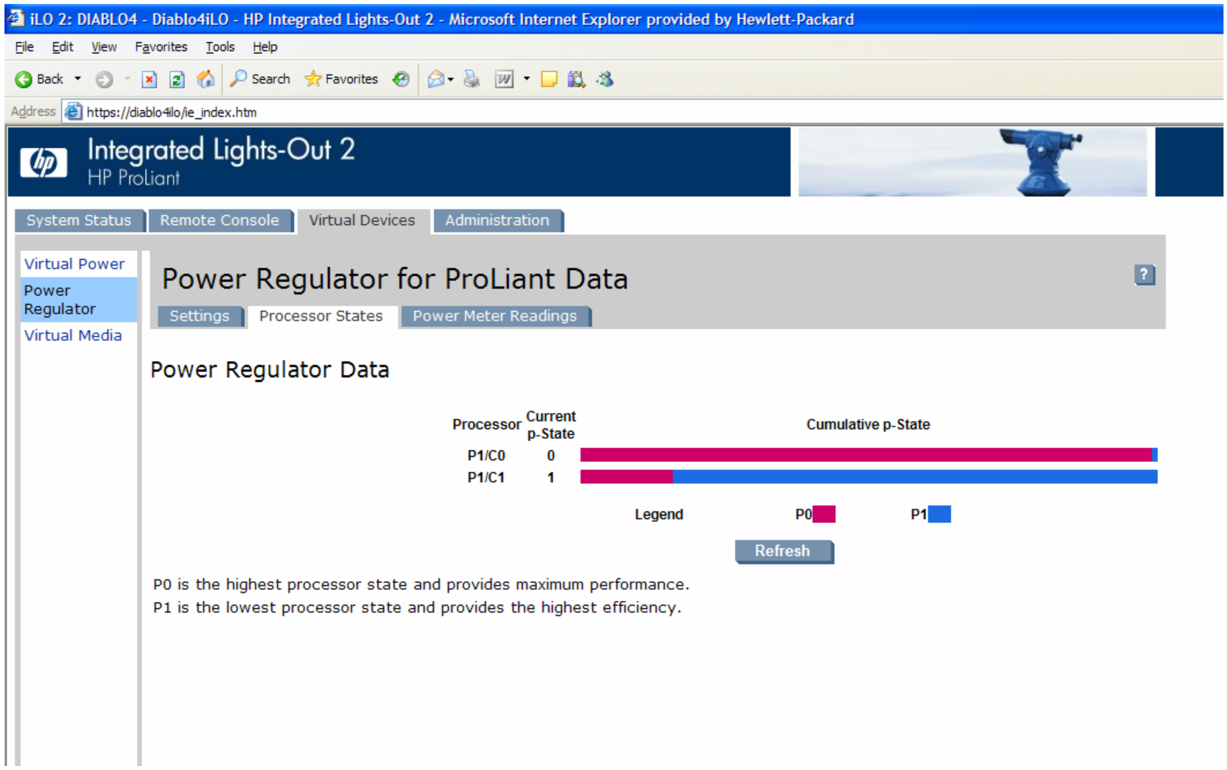
Power Regulator is an OS-independent power management feature of complaint HP ProLiant servers.[4] Power Regulator technology enables dynamic or static changes in CPU performance and power states. Power Regulator effectively executes automated policy-based power management at the individual server level. Power Regulator can be enabled on single systems from the system ROM, iLO browser, or command line. A more robust management capability is available by using HP SIM. HP SIM allows administrators to access the iLO management processors of multiple servers in a data center environment.

The iLO and iLO 2 management processors monitor the CPU operation of individual servers at high and low performance levels and report the percentage of time at each level over 24 hours. This provides an indication of CPU power efficiency. Results are accessible from the iLO browser, command line, and scripted reports (see Figure 4 for an example when using a browser). When IT administrators use SIM, this historical information is available for multiple servers along with more extensive reporting options accessed through HP Insight Power Manager (IPM). IPM allows administrators to view server power usage over time as well as over the entire data center, as described in more detail in the section titled "Efficient practices for facility-level power and cooling".

[3] Department of Energy: 8¢ per kWh US national average

[4] Information about Power Regulator compliant ProLiant servers can be found at; http://h18013.www1.hp.com/products/servers/management/ilo/sup_servers.html?jumpid=reg_R1002_USEN

Figure 4. Example of iLO Power Regulator reporting capability



Server virtualization using virtual machine technology

Several studies of the data center environment have shown that without server virtualization technology, typical x86 processor utilization rates range somewhere between 5 to 15 percent. IT administrators can use virtual machine technology (for example VMware, Microsoft Virtual Server, RHEL/XEN and SLES 10/Xen) to consolidate multiple older, physical server platforms onto a single more powerful and energy-efficient server platform. Therefore, virtualizing and consolidating servers can increase processor utilization rates, reduce capital expenses, and reduce operating expenses (such as physical space requirements, power, and cooling costs in the data center).

To learn more about virtualization, refer to the technology brief titled "Server virtualization" technologies for x86-based HP BladeSystem and HP ProLiant servers" referenced in the section titled "For more information" at the end of this paper.

Efficient practices at the rack level

Rack configuration affects the ability to power, cool, and control the servers in the rack. Air flow, air leakage, power redundancy, power distribution, and cable management all contribute to the fundamental issues in highly dense computing environments.

Rack configuration tools

Using configuration tools can assist data center managers in optimizing rack configurations for compatible hardware. The HP eCo-Enterprise Configurator provides factory default racking for HP hardware portfolio. This tool allows the user to build virtual server solutions online. With it, the user can choose servers, memory, operating system, storage, backup solution; configure power and cooling options; select appropriate software; and assign necessary services and support. This HP tool is available at: http://h30099.www3.hp.com/configurator/

It provides the height, weight, power, and thermal requirements for each system to be racked as well as the requirements for the fully configured rack. These specifications are crucial for data center managers who must know whether these specifications fall within acceptable parameters for their existing facility infrastructures, or for those managers who require planning data to support IT equipment upgrades.

Methods for calculating power supply requirements

Requirements for uninterruptible power supplies (UPS) can be calculated using sizers from the equipment maker. Using a UPS sizer is important to prevent over provisioning of power resources. A sizer for HP UPSes is available at: www.upssizer.com.

Power protection and management

UPS management modules can enable the administrator to monitor and manage power environments through comprehensive control of UPSes. Such modules can support either a single or multiple UPS configurations providing redundancy and no-single-point-of-failure.

The management module can be configured to send alert traps to networked remote management programs or be used as a stand-alone management system.

Rack based power distribution

The ability to monitor and manage power distribution units (PDU) is key to optimizing rack-based power distribution. PDUs can provide power displays for local monitoring and serial ports for remote monitoring. This can be augmented with a management module enabling the administrator to monitor power environments through comprehensive control of individual PDUs. Embedded network capabilities can also allow these modules to be tied into comprehensive remote management suites.

To learn more about HP examples of both UPS and PDU hardware and management tools, see the technology brief titled “Critical factors in intra-rack power distribution planning for high-density systems” in the “For more information” section at the end of this document.

High-efficiency power supplies

All ProLiant servers are equipped with high-efficiency switch-mode power supplies, when compared to typical power supplies in the industry. For example, a typical white-box server power supply has an efficiency rating between 65 percent and 70 percent. ProLiant servers operate with efficiencies of 85 percent or greater when connected to a high-line voltage source.

ProLiant server power supplies operate at maximum efficiency when connected to high-line input power (200 to 240 VAC). As with typical power supplies in the industry, operating at low line power (100 to 120 VAC) causes the power supply to operate at a lower efficiency and to draw more current for the same power output.

Understanding Internal airflow

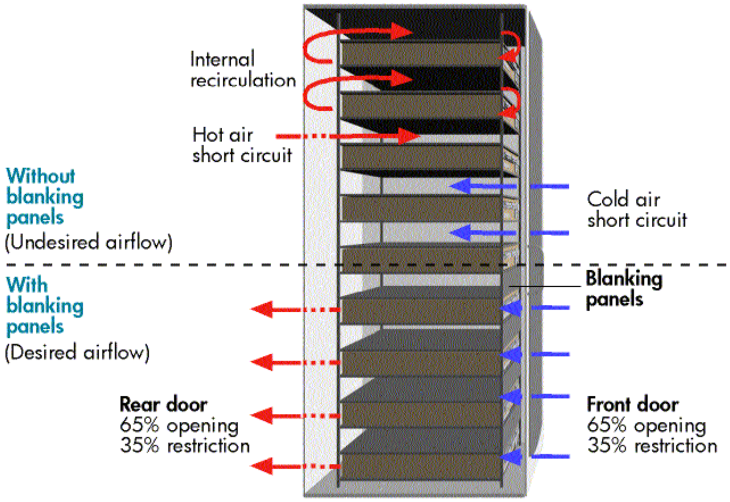
HP rack-mounted equipment is designed for front-to-back cooling. This means that the following conditions must exist:

- At least 65 percent of each of the front and rear rack surfaces must be open
- Airflow must be configured to force all air through the front and allow it to exhaust easily through the rear. This is promoted by using gaskets, blanking panels, and efficient cable management.

Even with the designation for cabinet doors that are 65 percent open to incoming airflow. There is still a 35 percent restriction to air discharged by the equipment in the rack. Servers will intake air from the path of least resistance. Therefore, they will access the higher-pressure discharge air flowing inside the cabinet mare easily than they will access cooling air coming through the front of the cabinet. Some configurations, such as those with extreme cable or server density, may create a backpressure situation forcing heated exhaust air around the side of a server and back into its inlet.

In addition, any gaps between the cold and hot isles allow hot exhaust air to re-circulate and mix with cold air, making a “short circuit”. This situation can occur when hot aisle exhaust air flows straight through a rack with open "U" spaces as shown in Figure 5. Gaskets or blanking panels must be installed in any open spaces in the front of the rack to support the front-to-back airflow design and prevent these negative effects. All rack space must be filled by equipment or enclosed by blanking panels so that the cool air is routed exclusively through the equipment and cannot bypass through or around the rack.

Figure 5. Airflow in rack without blanking panels (top) and with blanking panels (bottom)



Heat loads generated by blade servers in given rack configurations can be calculated by using the HP BladeSystem Power Sizer found at: www.hp.com/go/bladesystem/powercalculator.

Proper cable management practices:

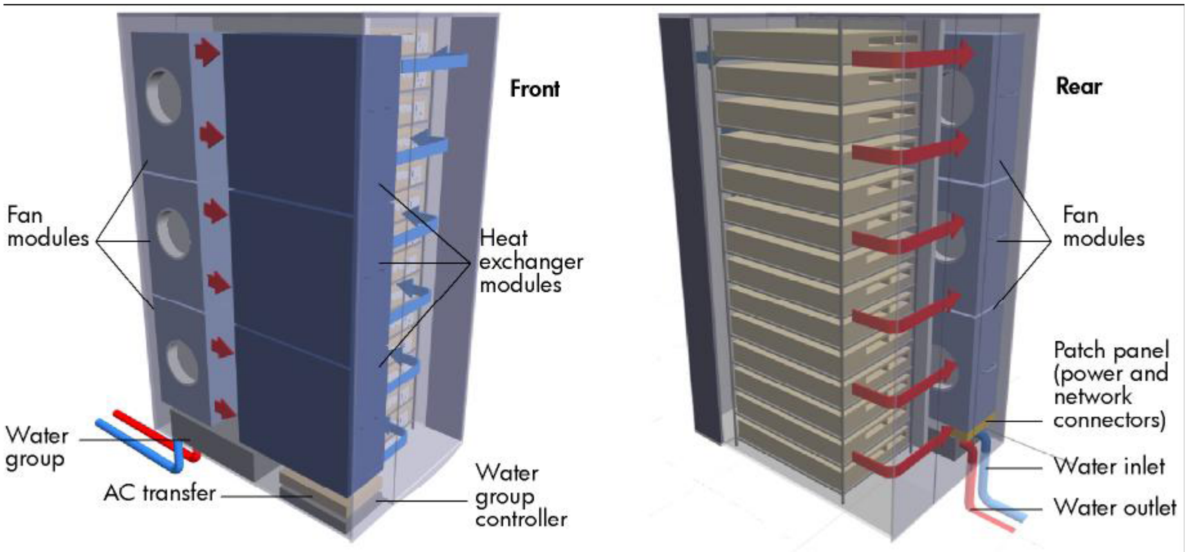
- Segregate power and data cables. Cables in close proximity to each other, especially those oriented in parallel or in loops, can create electromagnetic interference (EMI) through induction.
- Use cable routing and retention products. Cable rings, trays, and trunking products will reduce problems associated with cooling, EMI, and cable damage.
- Employ new cable trays to allow routing of cables above the racks rather than in plenum sub flooring, remove obstructions to airflow in the data center.
- Use CAT5 KVM cabling. The smaller, more flexible form factor of CAT5 cabling facilitates routing through the rack and its customizable lengths eliminate coils of cabling in the rear of the rack. The Interface Adapter (IA) that connects from the CAT5 cable to the server KVM port also provides keep-alive and naming functionality. Keep-alive means that even if the CAT5 cable is disconnected from the switch the IA will still respond as if the keyboard, mouse and video were present, allowing a server reboot to occur without errors.

Liquid cooled racks

The rack itself is not the limiting factor in high density rack configurations. Instead, the maximum heat load for air cooled racks is determined by the air volume and the air temperature being delivered to the rack. The average maximum load for air cooled racks is about 10kW/rack. The actual maximum load is dependant on the optimization of variables like data center design, placement of high density racks, the use of closely coupled cooling techniques, and the other best practices addressed in this document. When some or all of these techniques are optimized in concert, the actual load can go beyond 10 kW/rack. Liquid-cooled racks embody the concept of ”closely coupled” cooling where the cooling mechanism is in close proximity to the servers to be cooled and is aware of the specific

requirements of the target servers. Liquid-cooled racks are ideal for high density head loads. The HP Modular Cooling System (MCS), pictured in Figure 6, is one such rack. It is a closed-loop cooling system mounted on an HP 10000 Series G2 Rack. The MCS makes it possible to achieve hardware densities and power consumption levels (up to 35 kW in a single rack) that have been difficult—if not impossible—to support with conventional HVAC systems. In particular, the MCS allows data centers to resolve specific hot spot occurrences without revamping the overall infrastructure or introducing additional heat into the data center.

Figure 6. HP Modular Cooling System



An HP 10000 Series G2 rack with an attached MCS requires approximately 1½ times the width and 1¼ times the depth of a standard server rack (to allow for the fan and heat exchanger modules and front and rear airflow). However, one MCS enclosure has enough cooling capacity to support the heat load of a rack of equipment consuming 35 kW. This heat load is equivalent to that generated by three 10-kW racks, yet the MCS occupies 40 percent less floor space than three standard racks. Likewise, the MCS supports a heat load equivalent to 4.38 8kW racks (35 kW/8 kW per rack = 4.38 racks) while occupying 65 percent less floor space and reducing the overall heat load on the facility.[5]

Efficient practices for facility-level power and cooling

In the past when data centers mainly housed large mainframe computers, power and cooling design criteria were designated in average wattage per unit area (W/ft^2 or W/m^2) and British Thermal Units per hour (BTU/hr), respectively. These design criteria were based on the assumption that power and cooling requirements were uniform across the entire data center. Today, IT managers are populating data centers with a heterogeneous mix of high-density hardware as they try to extend the life of their existing space, making it important to understand power density distributions across the facility.

[5] Complete product information for the HP Modular Cooling System may be found at: http://h20000.www2.hp.com/bc/docs/support/SupportManual/c00600082/c00600082.pdf

As IT administrators and data center managers are well aware, facility power requirements involve much more than server power requirements. The percentage of total power consumption used by cooling alone in today's average data center can be as high as 70 percent.[6]

Power management and reporting

Management tools that provide graphing and historical analysis of key power and thermal data for servers in the data center can provide a comprehensive overview for facility metrics and management.

Comprehensive management applications

One tool with these attributes is HP's Insight Power Manager (IPM). Insight Power Manager is a ProLiant Essentials product that is part of HP SIM Graphing. Analysis is supported for single-server or multiple-server views. The IPM product provides graphing and historical analysis of key power and thermal data for supported ProLiant servers and it can store up to three years worth of power and thermal data. Since graphing and analysis is supported for single-server or multiple server views, administrators can estimate power and cooling costs for multiple servers, estimate peak consumption for multiple servers simultaneously, and compare data center temperatures across different parts of the data center.

Insight Power Manager also supports changing the Power Regulator mode for one or many ProLiant servers. Power Regulator changes may be made interactively via the web user interface or the changes may be scheduled to occur at specific and recurring times.

Power Capping

Using updated iLO 2 firmware (version 1.30) and updated System ROM/BIOS (dated 5/1/2007), selected HP ProLiant servers now have the ability to limit the amount of power consumed. Customers may set a limit in watts or Btu/hr. The purpose of this limit is to constrain the amount of power consumed, which reduces the heat output into the data center. The iLO 2 firmware monitors the power consumption of the server, checks it against the power cap goal, and, if necessary, adjusts the server's performance to maintain an average power consumption that is less than or equal to the power cap goal.

Using the IPM v1.10 plug-in to Systems Insight Manager v5.1, customers may set power caps on groups of supported servers. The IPM software statically allocates the group power cap among the servers in the group. The group cap is allocated equitably among all servers in the group based on a calculation using each server's idle and maximum measured power consumption.

The latest iLO 2 firmware may be found at http://www/hp.com/go/ilo. Updated System ROM/BIOS may be found on the Software and Drivers download page for each server model at www.hp.com/go/proliant. The latest Insight Power Manager software may be found at www.hp.com/go/ipm.

Airflow distribution for high-density data centers

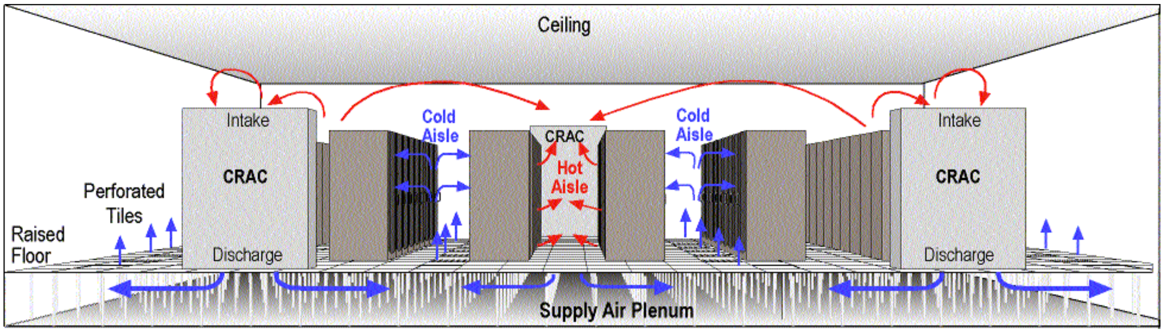
The front-to-rear airflow through HP equipment allows racks to be arranged in rows front-to-front and back-to-back to form alternating hot and cold aisles. The equipment draws in cold supply air from the front and exhausts warm air out the rear of the rack into hot aisles (Figure 7). Most data centers use a downdraft airflow pattern in which air currents are cooled and heated in a continuous convection cycle. The downdraft airflow pattern requires a raised floor configuration that forms an air supply

---

[6] Sources: Preliminary assessment from Uptime Institute: IDC Data Center of the Future US Server Power Spend for 2005 as a baseline($6bn); applied a cooling factor of 1; applied a .6 multiplier to US data for WW amount; Belady,C., Malone, C.,"Data Center Power Projection to 2014", 2006 ITHERM, San Diego, CA (June 2006)

plenum beneath the raised floor. The computer room air conditioning (CRAC) unit draws in warm air from the top, cools the air, and discharges it into the supply plenum beneath the floor.

Figure 7. Airflow pattern for raised floor configuration with hot aisles and cold aisles

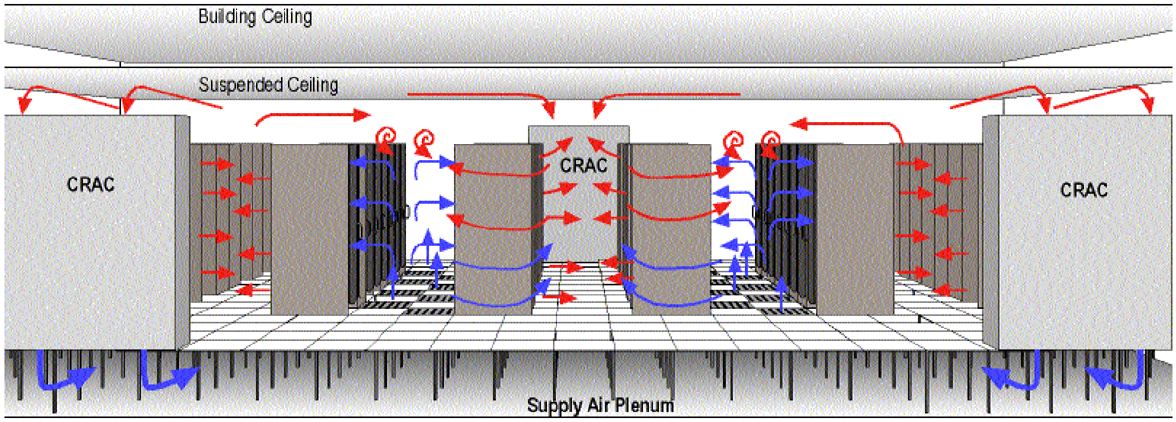


To achieve an optimum downdraft airflow pattern, warm exhaust air must be returned to the computer room air conditioning (CRAC) unit with minimal obstruction or redirection. Ideally, the warm exhaust air will rise to the ceiling and return to the CRAC unit intake. In reality, only the warm air closest to the intake may be captured; the rest may mix with the supply air. Mixing occurs if exhaust air goes into the cold aisles, if cold air goes into the hot aisles, or if there is insufficient ceiling height to allow for separation of the cold and warm air zones (Figure 8). When warm exhaust air mixes with supply air, two things can happen:

- The temperature of the exhaust air decreases, thereby lowering the useable capacity of the CRAC unit.
- The temperature of the supply increases, which causes warmer air to be re-circulated through computer equipment.

Therefore, administrators need to minimize the mixing of hot and cold air by using the practices outlined in the following sections.

Figure 8. Mixing of supply air and exhaust air



Raised floors

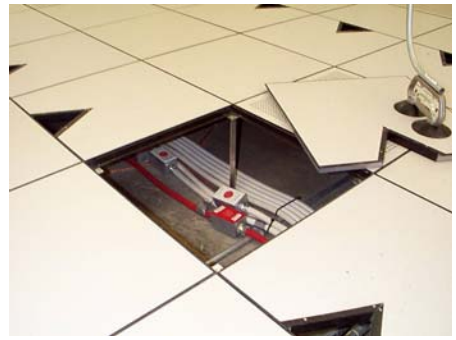
Raised floors typically measure 18 inches (46 cm) to 36 inches (91 cm) from the building floor to the top of the floor tiles, which are supported by a grounded grid structure. The static pressure in the supply plenum pushes the air up through perforated floor tiles to cool the racks. Most equipment draws in cold supply air from the front and exhausts warm air out the rear of the racks. Ideally, the warm exhaust air rises to the ceiling and returns along the ceiling back to the top of the CRAC units to repeat the cycle. Administrators should take into consideration that a higher supply air plenum (36 inches) will provide additional air flow for high density configurations.

Air supply plenum

The air supply plenum must be a totally enclosed space to achieve pressurization for efficient air distribution. The integrity of the subfloor perimeter (walls) is critical to prevent moisture retention and to maintain supply plenum pressure. This means that openings in the plenum perimeter and raised floor must be filled or sealed. Subfloor plenum dividers should be constructed in areas with large openings or with no subfloor perimeter walls.

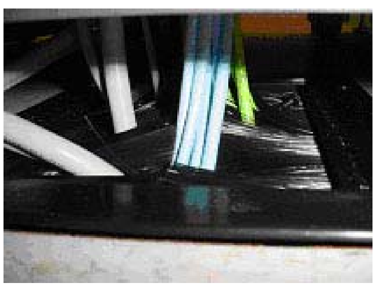
The plenum is also used to route piping, conduit, and cables that bring power and network connections to the racks. In some data centers, cables are simply laid on the floor in the plenum where they can become badly tangled (Figure 9). This can result in cable dams that block airflow or cause turbulence that minimizes airflow and creates hot spots above the floor. U-shaped “basket” cable trays or cable hangers can be used to manage cable paths, prevent blockage of airflow, and provide a path for future cable additions. Another option is to use overhead cable trays to route network and data cables so that only power cables remain in the floor plenum.

Figure 9. Unorganized cables (left) and organized cables (right) beneath a raised floor.



Electrical and network cables from devices in the racks pass through cutouts in the tile floor to wireways and cable trays beneath the floor. Oversized or unsealed cable cutouts allow supply air to escape from the plenum, thereby reducing the static pressure. Self-sealing cable cutouts are required to maintain the static pressure in the plenum (Figure 10). Cable management in high-density sever environments can significantly affect cooling issues.[7]

Figure 10. Self-sealing cable cutout in raised floor

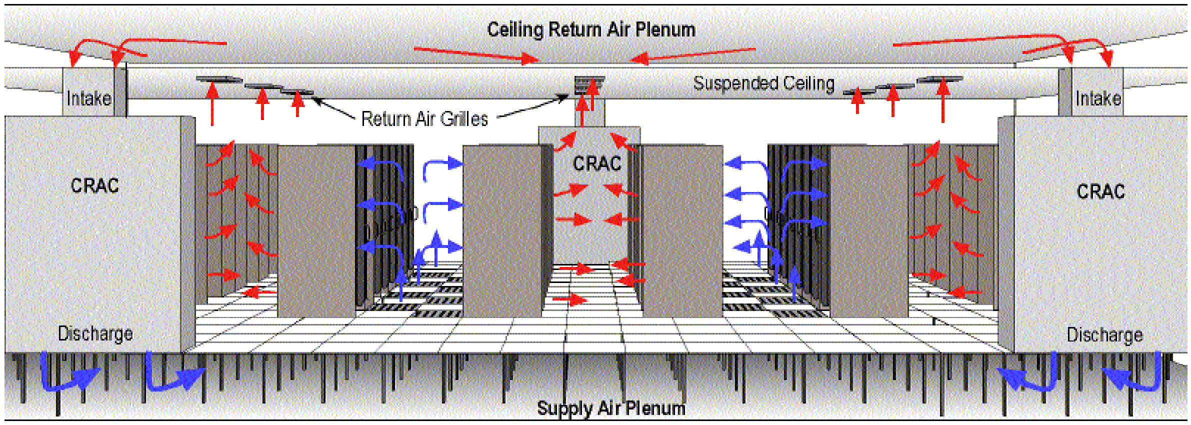


[7] For more information about cable management, see the technology brief titled "Cable management for rack-mounted systems," http://h20000.www2.hp.com/bc/docs/support/SupportManual/c01085208/c01085208.pdf

Ceiling return air plenum

In recent years, raised floor computer rooms with very high heat density loads have begun to use a ceiling return air plenum to direct exhaust air back to the CRAC intake. As shown on the right of Figure 11, the ceiling return air plenum removes heat while abating the mixing of cold air and exhaust air. Once the heated air is in the return air plenum, it can travel to the nearest CRAC unit intake. The return air grilles in the ceiling can be relocated if the layout of computer equipment changes.

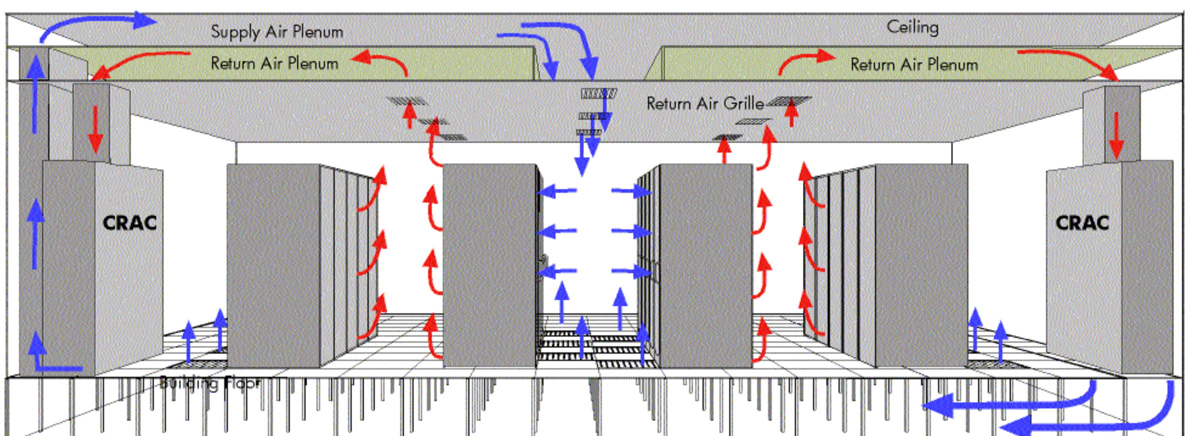
Figure 11. Ceiling return air plenum



Dual supply air plenums

As power and heat densities climb, a single supply air plenum under the raised floor may be insufficient to remove the heat that will be generated. High-density solutions may require dual supply air plenums, one above and one below (see Figure 12). In this configuration, additional supply air is forced downward in the cold aisle.

Figure 12. Dual air supply plenum configuration for high-density solutions



Perforated tiles

Floor tiles range from 18 inches (46 cm) to 24 inches (61 cm) square. The percentage and placement of perforated floor tiles are major factors in maintaining static pressure. Perforated tiles should be placed in front of at least every other rack. In higher density environments, perforated tiles may be necessary in the front of each rack. Perforated tiles are classified by their open area, which may vary

from 25 percent (the most common) to 56 percent (for high airflow). A 25 percent perforated tile provides approximately 500 cubic feet per minute (cfm) at a 5 percent static pressure drop, while a 56 percent perforated tile provides approximately 2000 cfm.

Rack geometry

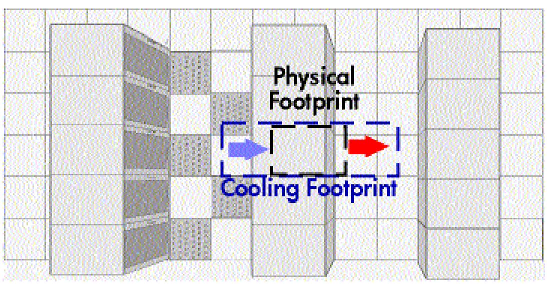
Designing the data center layout to form hot and cold aisles is one step in the cooling optimization process. Also critical is the geometry of the rack layout. Research by HP Laboratories has revealed that minor changes in rack placement can change the fluid mechanics inside a data center and lead to inefficient utilization of CRAC units. See the "Thermal Assessment Services" section for more information.

Cooling footprint

The floor area that each rack requires must include an unobstructed area to draw in and discharge air. Almost all HP equipment cools from front to rear so that it can be placed in racks positioned side-by-side. The cooling footprint (Figure 13) includes width and depth of the rack plus the area in front for drawing in cool air and the area in back for exhausting hot air.

Equipment that draws in air from the bottom or side or that exhausts air from the side or top will have a different cooling footprint. The total physical space required for the data center includes the cooling footprint of all the racks plus free space for aisles, ramps, and air distribution. Typically, a width of two floor tiles is needed in front of the rack, and a width of at least one unobstructed floor tile is needed behind the rack to facilitate cable routing.

Figure 13. Cooling footprint



Hot and cold aisle spacing

The amount of space between rows of racks is determined as follows.

- Cold aisle spacing should be 48 inches, two full tiles, and hot isle spacing should be at least one full tile, 24 inches minimum. This spacing is required for equipment installation and removal and for access beneath the floor.
- Cold aisles should be a minimum of 14 feet apart, center-to-center, or seven full tiles.

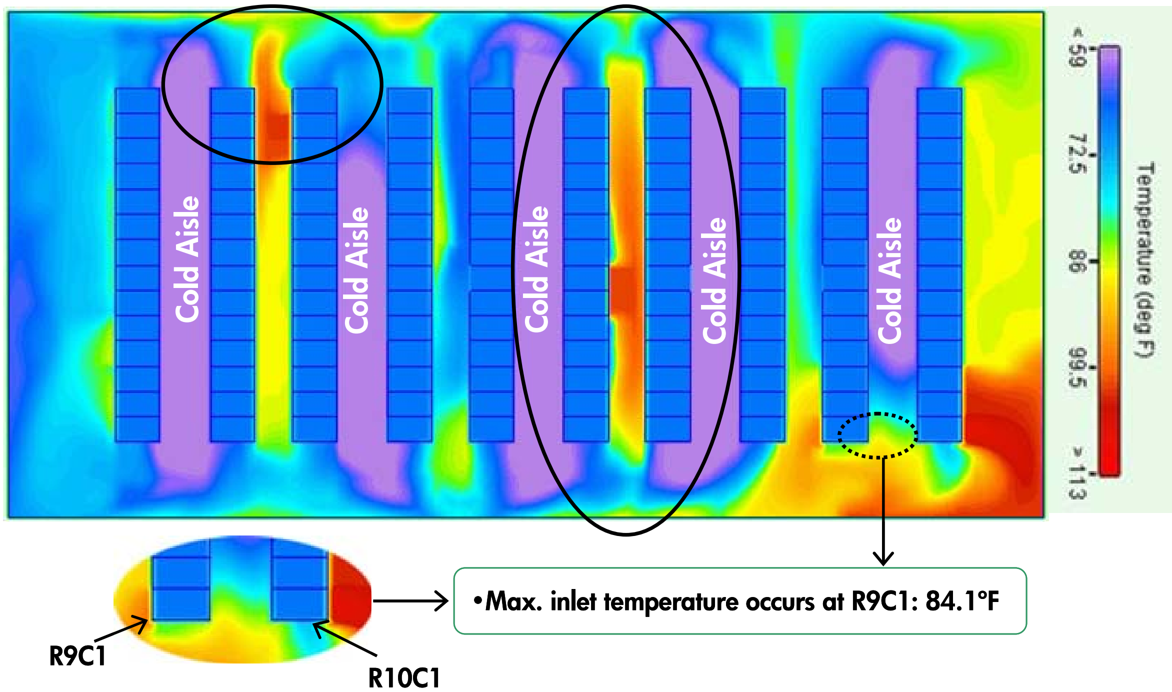
Row configuration

Keep equipment rows as long as safety requirements will allow and avoid row gaps to prevent mixing of the hot and cold air.

Where possible, locate high-density racks in the middle of equipment rows or mid-way between opposing air-conditioners. Avoid locating high-density racks at the ends of rows or deep into room corners.

At the ends of rows, exhaust from the high-density rack will wrap around the end of the row infiltrating the cold aisle and device inlets. Deep into room corners, heat will be trapped in corner and eventually increase inlet temperatures (see Figure 14).

Figure 14. Exhaust from the high-density rack wrapping around the end of the row



Closely coupled cooling

Cooling coupling is defined as how intimate and sensitive the cooling solution is to the individual IT equipment heat load. Close-coupling focuses on areas where the need for cooling is more pronounced, such as a rack populated with blade servers, rather than the heterogeneously-populated open space of the data center room. Close-coupling can result in shorter air paths that are more effective at heat extraction and require less fan power. Close-coupled heat removal minimizes the mixing of cool and hot air, since the airflow is completely contained in the row or rack.

Examples of closely coupled solutions are ceiling mounted heat exchangers, which target individual racks, and liquid cooled racks. Liquid cooled racks are particularly effective since they isolate the rack from the ambient conditions in the data center. There is no air mixing, and the cooling solution is immediately coupled with the rack which makes it highly energy efficient, as described in the section titled "Liquid cooled racks".

Computer room air conditioners

A common question with respect to cooling resources is how much capacity in kilowatts a particular CRAC unit can cool. Assuming a fixed heat load from the equipment in its airflow pattern, the answer depends largely on the capacity of the CRAC unit, its placement in the facility, and its discharge velocity.

Capacity of CRAC units

The heat load of equipment is normally specified in kilowatt per hour (kWh) or British Thermal Units per hour (BTU/hr). However, in the U.S., CRAC unit capacity is often expressed in "tons" of refrigeration, where one ton corresponds to a heat absorption rate of 3.5 kWh (12,000 BTU/hr).

While the "tons" capacity rating is measured at 80°F, HP recommends that facility managers target an optimal operating range of 68 - 77°F and 40 - 55% relative humidity (RH). These ranges are in inline with the recommendations of the American Society of Heating, Refrigerating and Air-Conditioning Engineers (ASHRAE)[8]. Managers should also be aware that as the operating temperature of the room decreases, so does the cooling capacity of the CRAC unit. Furthermore, the tons rating is very subjective because it is based on total cooling, which is comprised of "sensible cooling" and "latent cooling".[9]

Computer equipment produces sensible heat only; therefore, the sensible cooling capacity of a CRAC unit is the most useful value. For this reason, CRAC unit manufacturers typically provide cooling capacities as "total kWh" and "sensible kWh" (or "total BTU/hr" and "sensible BTU/hr") at various temperatures and RH values. Customers should review the manufacturer's specifications and then divide the sensible cooling capacity (at the desired operating temperature and humidity) by 3.5 kWh (12,000 BTU/hr) per ton to calculate the useable capacity of a given CRAC unit, expressed in tons of cooling.

Cooling capacity is also expressed in volume as cubic feet per minute (cfm). The volume of air required is related to the moisture content of the air and the temperature difference between the supply air and return air (ΔT):

Cubic feet per minute = (kWh x 3412) ÷ (1.08 × ΔT)

Or, using BTUs/hr

Cubic feet per minute = BTU/hr ÷ (1.08 × ΔT)

The cooling capacity calculations presented here are theoretical, so other factors must be considered to determine the effective range of a particular CRAC unit. The effective cooling range is determined by the capacity of the CRAC unit and the "cooling" load of the equipment in its airflow pattern. Typically, the most effective cooling begins about 8 feet (2.4 m) from the CRAC unit. The CRAC capacity, equipment cooling loads, and under floor conditions (airflow restrictions) will vary the effective cooling range of a CRAC.

NOTE:
Many CRAC manufacturers are now using kWh rather than BTU/hr to describe the capacities of their equipment. For that reason both units are included.

Placement of CRAC units

The geometry of the room and the heat load distribution of the equipment determine the best placement of the CRAC units. CRAC units can be placed inside or outside the data center walls. Customers should consider placing liquid-cooled units outside the data center to avoid damage to electrical equipment that could be caused by coolant leaks.

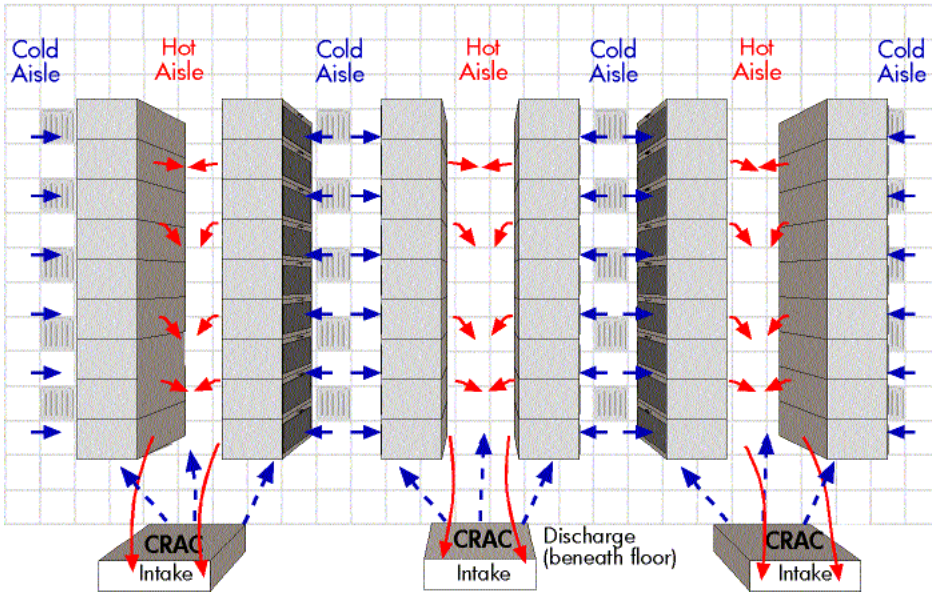
CRAC units should be placed perpendicular to the rows of equipment and aligned with the hot aisles, discharging air into the supply plenum in the same direction (Figure 16). This configuration provides the shortest possible distance for the hot air to return to the CRAC units. Discharging in the same direction eliminates dead zones that can occur beneath the floor when blowers oppose each other. Rooms that are long and narrow may be cooled effectively by placing CRAC units around the

[8] "Thermal Guidelines for Data Processing Environments" can be purchased at: http://resourcecenter.ashrae.org/store/ashrae/newstore.cgi?itemid=21074&view=item&page=1&loginid=2692065&words=thermal%20guidelines%20for%20data%20processing%20environments&method=and&

[9] Latent cooling is a result of a phase change (vapor and liquid) and sensible cooling is a result of lowering temperature

perimeter. Large, square rooms may require CRAC units to be placed around the perimeter and through the center of the room.

Figure 16. CRAC units should be placed perpendicular to hot aisles so that they discharge cool air beneath the floor in the same direction.



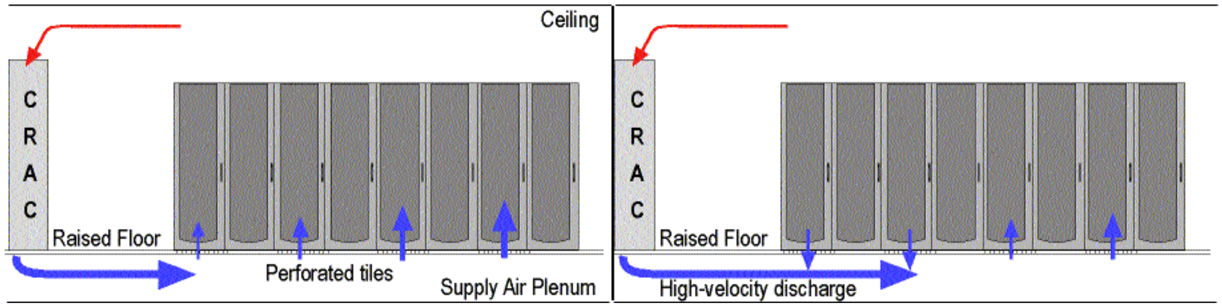
Discharge velocity

To force air from beneath the raised floor through the perforated tiles, the static pressure in the supply air plenum must be greater than the pressure above the raised floor. The velocity of the cooled air is highest near the CRAC unit because the entire flow is delivered through this area. The air velocity decreases as air flows through the perforated tiles away from the CRAC unit. The decrease in velocity is accompanied by an increase in static pressure with distance from the CRAC unit.

Excessive discharge velocity from the CRAC unit reduces the static pressure through perforated tiles nearest the unit, causing inadequate airflow (Figure 17). The static pressure increases as the high-velocity discharge moves away from the unit, thereby increasing the airflow through the perforated tiles. To counter this situation, airfoils under the raised floor can be used to divert air through the perforated tiles.[10] Another option is to use a fan-assisted perforated tile to increase the supply air circulation to a particular rack or hot spot. Fan-assisted tiles can provide 200 to 1500 cfm of supply air.

[10] From *Changing Cooling Requirements Leave Many Data Centers at Risk*. W. Pitt Turner IV, P.E. and Edward C. Koplin, P.E. ComputerSite Engineering, Inc.

Figure 17. Plenum static pressure greater than pressure above the floor (left). High-velocity discharge reduces static pressure closest to the unit (right).



Advanced thermal management techniques

Heat loads vary throughout a data center due to the heterogeneous mix of hardware types and models, changing compute workloads, and the addition or removal of racks over time. The variation in heat load may be too complex to predict intuitively or to address by adding cooling capacity. Approaches to managing these heterogeneous mixes of hardware and densities tend to be varied and proprietary. Hewlett Packard has been identified by independent consulting and research groups as a leader in advanced thermal management techniques for the data center. This section will focus on those HP techniques.

HP Laboratories has devised two thermal analysis approaches—Thermal assessment services[11] and Dynamic Smart Cooling—that manage heat distribution throughout a data center using computational fluid dynamics (CFD). Thermal assessment services use CFD modeling to aid planners in designing the physical layout of the data center for optimum distribution of cooling resources and heat loads. These modeling services can also predict the changes in heat extraction of each CRAC unit when the rack layout and equipment heat load are varied.

Dynamic Smart Cooling offers a higher level of automated facility management. It enables intelligent data centers that dynamically provision cooling resources to match the changing heat dissipation of computing, networking, and storage equipment. It also redistributes compute workloads based on the most efficient use of cooling resources within a data center or a global network of data centers.

Thermal Assessment Services

HP Thermal Assessment Services use CFD modeling to determine the best layout and provisioning of cooling resources based on fixed heat loads from data center equipment. The heat extraction of each CRAC unit is compared to its rated capacity to determine how efficiently (or inefficiently) the CRAC unit is being used, or "provisioned."

The provisioning of each unit in the data center is presented as a positive or negative percentage as follows:

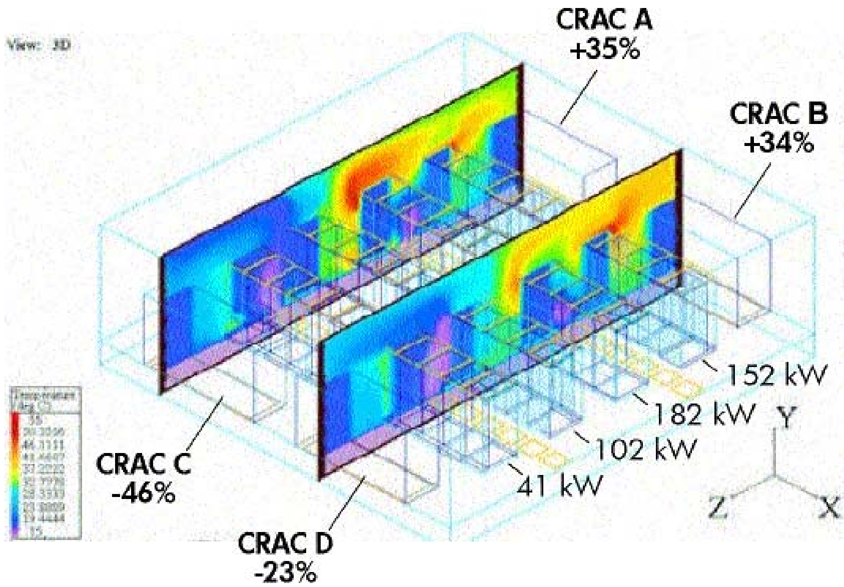
- An under-provisioned CRAC unit (positive percentage) indicates that the cooling load is higher than the capacity of the unit.
- A closely provisioned CRAC unit (small negative percentage) signifies that the cooling load is less than but reasonably close to the capacity of the unit, leading to efficient use of energy resources.

[11] For more information, please go to http://h20219.www2.hp.com/services/cache/114078-0-0-225-121.html?jumpid=reg_R1002_USEN.

- An over-provisioned CRAC unit (large negative percentage) operates significantly below the capacity of the unit. This results in wasted energy if operation of the unit cannot be adjusted to match the lower cooling load.

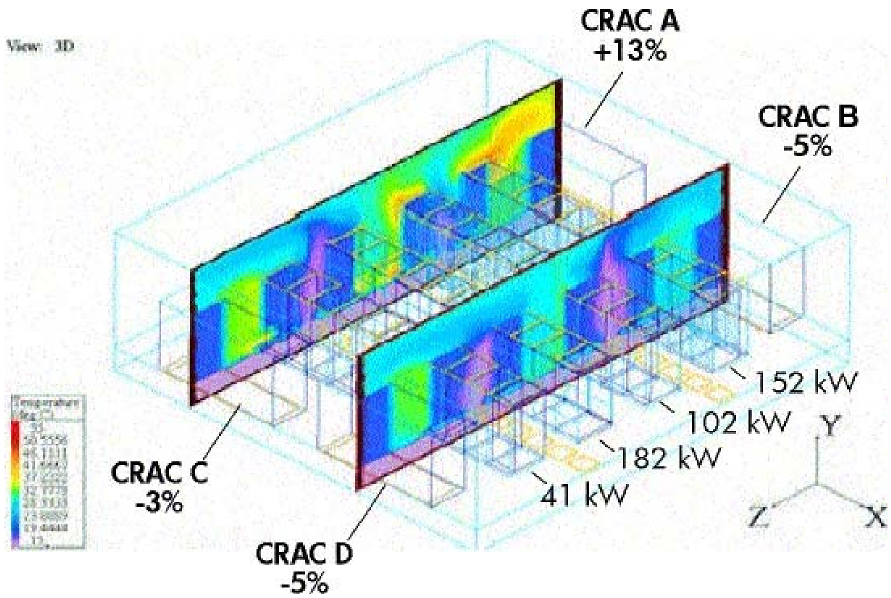
For example, Figure 18 shows the row-wise distribution of heat loads (41 kW to 182 kW) for a combination of compute, storage, and networking equipment in a typical raised floor data center with four CRAC units. The CFD model shows that the provisioning of the CRAC units is completely out of balance.

Figure 18. Poorly provisioned CRAC units



In Figure 19, the 102-kW row and the 182-kW row have been repositioned to better distribute the heat load. This CFD model shows that the CRAC units are now provisioned within 15 percent of their capacity.

Figure 19. Statically provisioned CRAC units



Dynamic Smart Cooling

Dynamic Smart Cooling (DSC) is a datacenter cooling solution which progresses the state-of-the-art beyond that which is feasible with any static solution. Static solutions lead to over-provisioning of resources due to the need to respond to peak demands, handle installation growth of the IT load and maintain sufficient levels of redundancy. Likewise, inadequate control leads to providing server inlet air at a temperature far below the required temperature, which leads to wasted capacity. Employing DSC in a data center requires a distributed monitoring system and a feedback control system that continually provisions the cooling resources based on the workload distribution.[12]

DSC is currently available for retrofit in existing datacenters or can be easily specified as part of new construction. It is important to emphasize that the intelligent control comprising DSC is not a replacement for best practices in the datacenter. Rather, employing DSC in conjunction with the guidance provided in this paper will maximize potential benefits in the datacenter. For more information on DSC, consult the "Data Center Cooling Strategies" technical brief available at http://h20000.www2.hp.com/bc/docs/support/SupportManual/c01153741/c01153741.pdf.

[12] Patel, C.D., Sharma, R.K, Bash, C.E., Beitelmal, A, Friedrich, R., "Smart Cooling of Data Centers," July 2003, IPACK2003-35059, Proceedings of IPACK03- International Electronics Packaging Technical Conference and Exhibition, Maui, Hawaii.

Summary

Data centers are approaching the point of outpacing conventional methods used to power and cool high density computing environments. Escalating energy costs and cooling requirements in existing data center facilities call for better methodology in the areas of planning and configuration, and more capable analytical and management tools to handle power and cooling demands.

Data center and facility managers can use best practices to greatly reduce the heat loads. These practices include:

- View data center and data center components as a completely integrated infrastructure
- Assess existing facility power and cooling resources
- Maximize power and cooling capabilities at the component level
- Optimize facility for efficient power distribution
- Institute floor plans and standard practices that maximize rack and aisle cooling
- Promote highly automated and virtualized data center operation
- Manage power and cooling as variable resources that dynamically respond to processing
- Employ continuous and comprehensive monitoring
- Choose an integrated approach to data center hardware, software, applications, network and facility

To address more complex situations HP Professional Services can work directly with customers to optimize existing data centers for more efficient cooling and energy consumption. The Thermal Assessment services can also be used to confirm new data center designs or predict what will happen in a room when certain equipment fails. As long as the data center has the power and cooling resources to support the expected loads, Thermal Assessment Services can rectify cooling problems as well as enhance the overall efficiency of air conditioning resources. In most cases, the energy savings alone may pay for the cost of the service in a relatively short period.

Such modeling services supply data center managers with more accurate mapping of air flow and temperature distribution in the facility. This in conjunction with solid fundamentals such as closely coupled cooling, rack geometry, cooling footprints, and rack air flow management provide a sound basis for effective power and cooling management.

Dynamic Smart Cooling (DSC) is the result of HP taking a 'holistic' view of data center management, analyzing the requirements, and building an adaptive control system around that analysis. The resulting DSC technology will enable managers to prevent over-provisioning of cooling and power resources and provide the adaptive control to direct those resources where needed.

For more information

For additional information, refer to the resources detailed below.

Resource description	Web address
Thermal Considerations in Cooling Large Scale High Compute Density Data Centers white paper	www.hpl.hp.com/research/papers/2002/thermal_may02.pdf
HP Rack/Site Installation Preparation Utility	http://h30099.www3.hp.com/configurator/calc/Site%20Preparation%20Utility.xls
Power calculators	http://h30099.www3.hp.com/configurator/calc/Power%20Calculator%20Catalog.xls
C7000 Blade enclosure	http://h20000.www2.hp.com/bc/docs/support/SupportManual/c00816246/c00816246.pdf
Cable management for rack-mounted systems	http://h20000.www2.hp.com/bc/docs/support/SupportManual/c01085208/c01085208.pdf
HP Modular Cooling System	http://h20000.www2.hp.com/bc/docs/support/SupportManual/c00600082/c00600082.pdf
Critical factors in intra-rack power distribution planning for high-density systems	http://h20000.www2.hp.com/bc/docs/support/SupportManual/c01034757/c01034757.pdf
Server virtualization technologies for x86-based HP BladeSystem and HP ProLiant servers	http://h20000.www2.hp.com/bc/docs/support/SupportManual/c01067846/c01067846.pdf
Data Center Cooling Strategies	http://h20000.www2.hp.com/bc/docs/support/SupportManual/c01153741/c01153741.pdf

Call to action

Send comments about this paper to TechCom@HP.com.

TC070802TB, August 2007